Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Coherent Multidimensional Poverty Measurement

Gaël GIRAUD

2012.95

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/bandeau-haut/documents-de-travail/
ISSN : 1955-611X

Coherent Multidimensional Poverty Measurement*

Gaël Giraud†
CNRS, CES, Paris School of Economics

January 10, 2013

Abstract

This paper presents a family of multidimensional poverty indices that measure poverty as a function of the extent and the intensity of poverty. I provide a unique axiomatics from which both extent and intensity of poverty can be derived, as well as the poor be endogenously identified. This axiomatics gives rise to a family of multidimensional indices whose extremal points are the geometric mean and the Maximin solution. I show that, in addition to all the standard features studied in the literature, these indices are continuous (a must for cardinal poverty measures) *and* ordinal, in the sense that they do not depend upon the units in which dimensions of achievements are computed. Moreover, they verify the decreasing rate marginal substitution property: the higher one's deprivation (or the extent of poverty) in one dimension, the smaller the increase of achievement in that dimension that suffices to compensate for a decrease of achievement in another dimension.

Keywords: multidimensional poverty, geometric mean, Maximin solution, utilitarian solution, endogenous identification, coherence, continuity, decreasing marginal rate of substitution, cardinal date, ordinality, relative weights
JEL Classification Numbers : I3, I32, D31, D63, O1

1 Introduction

As acknowledged by Villar (2010), "defining a poverty measure in a truly multidimensional context involves a number of subtle and difficult issues: choosing the appropriate poverty dimensions beyond income or wealth, deciding on whether they all are equally important, fixing sensible thresholds in those dimensions and setting criteria to identify as poor those individuals whose achievements lie partially below them, defining an overall measure of poverty intensity, etc. Those difficulties

*I thank Cécile Renouard, as well as Camille Sutter, Raphaële de la Martinière, Suman Seth and participants of the HDCA 2012 congress (Jakarta) for fruitful discussions.
†gael.giraud@parisschoolofeconomics.fr

anticipate that many compromises are required and, indirectly, that the axiomatic approach may be the best way to deal with this type of problem as it makes explicit all those compromises." [1]

Here, we provide an axiomatization for a family of Multidimensional Poverty Indices. This is part of a larger research programme devoted to a Relational Capabilty Index applied as a new poverty measure in Nigeria, Indonesia and India.

Each index can be characterized as lying somewhere between the two extremal points of our family of indices: the geometric mean (Villar (2010)) and the Rawlsian Maximin. Although both social choice correspondences have been thoroughly studied from the social choice theoretical viewpoint,[2] we are not aware of any attempt to link these two major concepts of justice with the concerns involved in the literature devoted to poverty measurement. This paper is a first attempt to fill this gap.

We suggest that the geometric mean can be interpreted as being a (hyperbolic) version of the "utilitarian" viewpoint. With this interpretation in mind, our family of indices builds a bridge between celebrated theories of justice and poverty measurements. An alternative standpoint enables us to characterize each one of our indices as being the supremum of the weighted geometric averages, the sup being taken over some collection of weights over dimensions and people. When the collection of weights reduces to the uniform vector, we are back to the standard geometric mean (this is the "utilitarian" solution). When the underlying collection of weights includes the whole unit simplex over dimensions and people, then we get the Maximin solution. One possible interpretation is as follows: suppose that the economist who is in charge of measuring poverty in a given population reflects as if she were in Rawls' original position. Beyond the veil of ignorance, the point that is ignored is not which role one will endorse (as in the standard, political interpretation of Rawl's theory of justice) but in which dimension one will get some talent (or some endowment, or some "social capital"). So uncertainty bears on dimensions rather than on persons. In addition, from the viewpoint of the economist standing beyond the veil, there might be some ambiguity concerning the probability according to which talents and deprivations will be distributed. As a result, if the economist has no prejudice about the distribution of talents and deprivations will be distributed, she might opt for the Maximin solution as a way to measure multidimensional poverty. if, on the contrary, she has good reasons to believe that the distribution will be uniform, she may choose the "utilitarian"' solution (i.e., in our context, the geometric average). Else, she might choose an index in our family which lies somewhere in between the first two. If one wishes so, it is also possible to include ambiguity about the persons (and not only dimensions) in the non-symmetric version of our family of indices.

To the best of our knowledge, this is the fist attempt to formally fill the gap between theories of justice and poverty measurements. As we take inspiration from Artzner et al. (1999) (where an additive version of a similar axiomatization was introduced in order to measure the risk position of a portfolio), we call *coherent* a multidimensional poverty index belonging to our family. We finally show that such indices satisfy the following properties that are considered as desirable in the literature:[3]

[1]See Dardadoni (1995), Ravallion (1996), Tsui (2002), Bourguignon and Chakravarty (2003), Lugo and Maasoumi (2008), Alkire and Foster (2008), Wagle (2008), and **?**.

[2]See **?** and **?** to name but a few pathbreaking papers in this area.

[3]See, e.g., Bourguignon and Chakravarty (2003) and Alkire and Foster (2008).

(i) Each index is continuous: slight changes in the achievements of certain persons only induces slight changes in the poverty measurement ;

(ii) Each index is ordinal, in the sense that it does not depend upon the choice of the specific units in which dimensions of achievements are measured. This property deserves some comment. In Alkire and Foster (2008) it is argued that data describing capabilities and functionings in the spirit of Sen's Multidimensional Human Index, are ordinal in nature. They therefore may lack a basis for comparisons across dimensions. This, of course, is a challenge for Mutldimensional Poverty measurement. In the above quoted paper, indeed, only one kind of measures is shows to be ordinal in that sense (the M^0 measure in their parlance) while they others don't. At the same time, this ordinal measure fails to satisfy a number of other properties. In particular, it cannot capture the intensity of poverty —a failure that can be viewed as arising from its being a piecewise constant (hence discontinuous) measure. Here, we prove that coherent Poverty measures are ordinal in the following sense: If one multiplies any dimension by $\lambda > 0$ (both for achievements and for the poverty cut-off), then the set of poor is unaffected while the Index, P, is simply multiplied by λ. As a consequence, a normalized version of the index, Q, is independent of such changes in the dimensions' unit.

(iii) it yields a criterion for "relative poverty" that depends upon the whole population under scrutiny ;

(iv) the marginal rate of substitution among subjects or among dimensions is decreasing. The reduction in the deprivation[4] of dimension k for poor individual i required to compensate an increase in the deprivation of dimension k for individual h is larger the higher the initial level of deprivation in i.

(v) As in Villar (2010), it is multiplicatively decomposable by population subgroups (but it does not satisfy Subgroup Decomposability in the additive form given in Bourguignon and Chakravarty (2003)). This property says the following: If the population is partitioned into subgroups, the overall poverty index corresponds to the weighted average of subgroup poverty values, where the weights correspond to population shares.

(vi) In certain circumstances, we may have additional informations that allow us to regard certain dimensions and/or certain subgroups of the population as meriting a greater relative weight than others. Each index can be adjusted so as to capture this kind of requirements. Of course, if one wishes so, it can as well be made symmetric among persons.

(vii) It verifies the transfer principle (Villar (2010)) : a reduction of size $\delta > 0$ in the deprivation with respect to dimension k of a poor person i who is worse off in this dimension than another poor person, j, more than compensates an increase of the same size, δ, in the deprivation of j, provided their relative positions remain unaltered.

(viii) Principle of population: a replica of the population does not change the poverty measure.

To the best of our knowledge, coherent poverty measures are the first examples of continuous and ordinal Multidimensional Poverty measure that are sensitive to inequality. To take but alternative examples, the measure M^0 introduced in **?** is ordinal but discontinuous and inequality-insensitive. On the other hand, the measures

[4]Recall that a person is said to be *deprived* in one dimension whenever her achievement falls below the cut-off or dimension-specific poverty line.

M^1 and M^2 are inequality-sensitive and continuous but no more ordinal.

The paper is organized as follows. The next section provides the model and makes the link between the "utilitarian" standpoint and the geometric mean explicit. Section 3 deals with the axiomatization of coherent multidimensional poverty indices. The last section provides the main properties of this family of indices.

2 The model

Let $\mathbf{N} = \{1, ..., N\}$ denote a society consisting of n individuals and let $\mathbf{K} = \{1, ..., K\}$ be a set of dimensions.

A social state is a matrix, $y = \left(y_{ij}\right)_{ij} \in \mathcal{M}_{N\times K}(\mathbb{R}_{++})$, with N rows, one for each individual, and K columns, one for each dimension. The entry $y_{ij} \in \mathbb{R}_{++}$ describes the value of variable j for individual i. Since we are going to deal with ordinal Poverty measures, there is little loss of generality in imposing from the outset that all variables be strictly positive.[5] A vector $z \in \mathbb{R}^K_{++}$ of reference values describes the poverty thresholds for all dimensions. How these thresholds are fixed is definitely an important issue, but we leave it aside here. We denote by $N(y; \mathbf{z}) \subset \mathbf{N}$ the set of poor that results from a social state matrix y and a vector z of reference values. The number of poor people is $n(y; \mathbf{z}) := |N(y; \mathbf{z})|$. As we shall see, $N(y; \mathbf{z})$ (hence $n(y; \mathbf{z})$) will be determined endogenously by our multidimensional poverty index).

2.1 The utilitarian index

A poverty index is a mapping $P : \mathbb{R}^{KN}_{++} \to \mathbb{R}_+$. We begin with three axioms that unambiguously characterize the "utilitarian" Poverty index.

The first one, anonymity, says that all agents and all dimensions are equally important:

> Anonymity. Let $x \in \mathbb{R}^{KN}_{++}$ and let $\pi \in \mathcal{S}_{KN}$ denote a permutation over its components $\{1, ..., kn\}$. Then, $P(s) = P(\pi(s))$.

The second Axiom implies that P reduces to the identity mapping on the diagonal of $\mathbb{R}^{KN}_{+}$:

> Normalization. Let $s \in \mathbb{R}^{KN}_{++}$ be such that $s_{ij} = a \; \forall i, j$. Then $P(s) = a$.

The last property requires that the difference between the new and the initial values of P when subject i's achievement relative to dimension j changes from s_{ij} to t_{ij}, be a monotone function of the difference between s_{ij} and t_{ij}.

[5] We cannot claim that this entails no loss of generality at all. Indeed, if some achievement is "naturally" given as being (strictly) positive, then, whether it is scored $\varepsilon > 0$ or $\lambda\varepsilon > 0$ does not matter. However, if the original achievement was 0, then, replacing it by $\varepsilon > 0$ might have an effect on the poverty measure. The same problem arises, e.g., in Seth (2009).

Difference Monotonicity Let $s, t \in \mathbb{R}_+^{KN}$ be such that $s_{hq} = t_{hq}$ $\forall (h, q) \neq (i, j)$. Then

$$P(s) - P(t) = g_{ij}(s_{ij} - t_{ij}),$$

for some increasing function $g_{ij} : \mathbb{R}_+ \to \mathbb{R}$. Since $g_{ij}(0) = 0$, it follows that $g_{ij}(x) \geq 0$ if, and only if, $x \geq 0$.

Proposition 2.1 *An index $P(\cdot)$ satisfies Anonymity, Normalization and Difference Monotonicity if, and only if, it takes the form*

$$P(s) = \frac{1}{kn} \sum_{i \in N, j \in K} s_{ij}.$$

Proof. Let $s \in \mathbb{R}_+^{KN}$. By difference monotonicity and normalization,

$$\begin{aligned} P(s_{11}, 0, ..., 0) - P(0, ..., 0) &= g_{11}(s_{11}) \\ P(s_{11}, s_{12}, ..., 0) - P(s_{11}, 0, ..., 0) &= g_{12}(s_{12}) \end{aligned}$$

so that

$$P(s) = P(0) + \sum_{i,j} g_{ij}(s_{ij}).$$

By anonymity, $g_{ij}(\cdot) = g(\cdot)$ $\forall i, j$. The Normalization axiom yields: $P(0, ..., 0) = 0$. Moreover,

$$P(a, ..., a) = kng(a) = a.$$

Therefore, $g(a) = \frac{a}{kn}$. The conclusion follows.

□

2.2 The geometric average

The link between the (admittedly fairly classical) index, P_U, and the geometric average is given by the following transformation :

Consider the following Poverty index, $G(\cdot)$, defined on $\mathbb{R}_{++}^{KN}$:

$$G(x) := \left[\prod_{k,h} x_{k,h} \right]^{\frac{1}{kn}}.$$

Given a vector, $x \in \mathbb{R}_{++}^{KN}$, let us denote by $\ln x$ the vector whose coordinates are $\ln x_{k,h}$, every h, k. Obviously,

$$G(x) = \exp U(\ln x). \tag{1}$$

From this very simple remark, one deduces the axiomatization provided by Villar (2010) that fully characterizes the geometric average as a Poverty index: Indeed, it follows from (1) that G must verify the anonymity and normalization Axioms together with the following ratio monotonicity:

Ratio Monotonicity Let $s, t \in \mathbb{R}^{KN}_{++}$ be such that $s_{hq} = t_{hq}$ $\forall (h,q) \neq (i,j)$. Then,

$$\frac{G(s)}{G(t)} = g_{ij}\Big(\frac{s_{ij}}{t_{ij}}\Big),$$

for some increasing function $g_{ij} : \mathbb{R}_{++} \to \mathbb{R}$. Since $g_{ij}(1) = 1$, it follows that $g_{ij}(x) \geq 1$ if, and only if, $x \geq 1$.

In other words, the geometric (or Cobb-Douglas) average may be viewed as the outcome of the Utilitarian rule after the transformation given by (1). In the following section, we show that $G(\cdot)$ is but one extremal point of a whole family of Poverty indices that can be constructed in quite a similar way. The other extremal index of this family turns out to be the Maximin rule.

3 Coherent Poverty Indices

In order to define a coherent Poverty index, we impose some axioms on the mapping $P(\cdot)$. For this purpose, we define a *poverty exit set*, $\mathcal{E} \subset \mathbb{R}^{KN}_{++}$. A population belongs to $\mathcal{E}$, whenever it is *not* poor.

3.1 Axioms for $\mathcal{E}$.

We take $\mathbf{z} \in \mathbb{R}^{KN}_{++}$ as given. The vector $\mathbf{z}$ describes the poverty thresholds for all dimensions and every individual. Those reference values may be fixed externally (absolute poverty lines) or may depend on the data of the social state matrix itself (relative poverty lines, such as a fraction of the median or the mean value). The units in which achievements are measured are chosen so that $\mathbf{z} >> \mathbf{1} := (1, ..., 1) \in \mathbb{R}^{KN}_{++}$.[6] When all individuals are attribtued the same cut-offs, $\mathbf{z} = (z, ..., z)$, for some reference vector $z \in \mathbb{R}^{k}_{++}$. In this case, if $x_i \geq z$, then *individual* i can be said to be poor (the converse being false in general).

In order to build an *ordinal* index (i.e., an index that does not depend upon the choice of unities in which dimensions are measured), we consider only normalized achievements. That is, if $x \in \mathbb{R}^{KN}_{++}$ is a given achievement, we shall deal with $\mathbf{x} := (x_{hk}/\mathbf{z}_{hk})_{h,k}$.

Axiom 1. $\mathbf{1} + \mathbb{R}^{KN}_{+} \subset \mathcal{E}$.

Consider, now, a population such that *all* its individuals have achievements (before normalization) that are *all* strictly below the thresholds given by $\mathbf{z}$. Obviously, such a population should be considered as poor. This is the content of the next Axiom.[7]

[6] Given two vectors x, y, $x << y$ if the strict inequality holds coordinatewise.

[7] Actually, Axiom 2 says a little bit more since it implies that, in the case of a single person population, this person will be poor if none of her achievements end up strictly above the cut-off, and at least one of her achievements stay strictly below this cut-off. We could replace Axiom 2 by the weaker $\mathcal{E} \cap \left(\mathbf{z} + \mathbb{R}^{KN}_{--}\right) = \emptyset$. But our other axioms would nevertheless strengthen it into Axiom 2 in most cases of interest for practical purposes.

Axiom 2. $\mathcal{E} \cap \left(\mathbf{1} + \mathbb{R}_{-}^{KN}\right) = \{\mathbf{1}\}$.

For every $x, y \in \mathbb{R}_{++}^{KN}$, let $x \Box z$ denote the vector in $\mathbb{R}_{++}^{KN}$ whose coordinates are $x_{k,h} z_{k,h}$, all k, h. Consequently, $1/y$ denotes the (unique) vector such that $y \Box 1/y = \mathbf{1}$, while x^λ is the vector with coordinates $x_{k,h}^\lambda$. The "box" product, $\cdot \Box \cdot$, is to be interpreted as formalizing a change in the achievements of the population under scrutiny. For instance, $x \Box \mathbf{1} = x$ stands for "no change". By contrast, $x \Box 0 = 0$ represents a radical depletion of the population, etc. For an arbitrary vector, $y \in \mathbb{R}_{++}^{KN}$, $x \Box y$ will represent a change that may be dimension- and individual-dependent.
A set $\mathcal{F} \subset \mathbb{R}_{++}^{KN}$ is *multiplicatively convex* whenever, as soon as $x, y \in \mathcal{F}$, then $x^\alpha \Box y^{1-\alpha} \in \mathcal{F}\ \forall \alpha \in [0,1]$.

Axiom 3. The Poverty exit set, $\mathcal{E}$, is multiplicatively convex

Axiom 1 says that, if *all* the individuals of a population exhibit all their achievements above the threshold (i.e., if $\mathbf{x} \geq \mathbf{1}$), this population is not poor. Conversely, if $\mathbf{x} << 1$, Axiom 2 implies that the population is poor. Ambiguity remains only whenever *some* individuals exhibit *some* achievements above the threshold, and others, not. Our last axiom deals with such ambiguous cases. Suppose that a population, x, is not poor. Take $\lambda > 0$ and consider the auxiliary population given by x^λ. Axiom 4 says that this new population should not be considered as poor neither. Clearly, if $x \Box 1/\mathbf{z} \geq \mathbf{1}$ (resp. $x \Box 1/\mathbf{z} < \mathbf{1}$), then $(x^\lambda \Box 1/\mathbf{z}) \geq \mathbf{1}$ (resp. $< \mathbf{1}$), so that the auxiliary population turns out, indeed, not to be poor (resp. to be poor). What the next Axiom says is that this property should not hold only for the extreme cases envisaged by Axioms 1 and 2 but also for the "ambiguous" cases.
A set $\mathcal{F} \subset \mathbb{R}_{++}^{KN}$ is a multiplicative cone whenever, as soon as $x \in \mathcal{F}$, then $x^\lambda \in \mathcal{F}$ for any $\lambda \geq 0$.

Axiom 4. The Poverty exit set, $\mathcal{E}$, is a multiplicative cone.

Examples The two following sets verify all four axioms.
a) The "utilitarian case". Consider

$$\mathcal{E} := \{\mathbf{x} \in \mathbb{R}_{++}^{KN} \mid G(\mathbf{x}) \geq G(\mathbf{z})\}.$$

$\mathcal{E}$ is the upper-set of the hyperbola $\{\mathbf{x} \ : \ G(\mathbf{x}) = \lambda\}$, and is closed and (additively) convex.
b) The "Rawlsian case". Consider

$$\mathcal{E} := \{\mathbf{x} \geq \mathbf{1}\},$$

$\mathcal{E}$ is an affine closed, convex (additive) cone.

Although it is not necessary for the core of our theory, the next Axiom will prove handful.

Axiom 5. $\mathcal{E}$ verifies the following Anonymity property : Let $x = (x_1, .., x_n) \in \mathbb{R}_{++}^{KN}$ and $\sigma(x) = (x_{\sigma(1)}, ..., x_{\sigma(N)}) \in \mathbb{R}_{++}^{KN}$ the vector obtained after having swapped its individuals with the permutation $\sigma \in \mathcal{S}_N$. Then,

$$\mathbf{x} \in \mathcal{E} \iff \sigma(\mathbf{x}) \in \mathcal{E}\ \forall \sigma \in \mathcal{S}_N.$$

3.2 Axioms for P

Given a set $\mathcal{E}$, the Poverty index, $P_{\mathcal{E}}$, is defined as a measure of the minimal additional "achievements" that have to be added to a given distribution so that the population can be considered as non-poor, i.e., so that the resulting distribution belongs to $\mathcal{E}$. Obviously, $P_{\mathcal{E}}$ will heavily depend upon the choice of $\mathcal{E}$. This is why the axiomatic studied here only defines a *family* of poverty indices, each index being identified through its corresponding poverty exit set.

Given $\mathcal{E}$, the mapping $P_{\mathcal{E}}(\cdot)$ is defined as:

$$P_{\mathcal{E}}(\mathbf{x}) := \inf\big\{\tau \in \mathbb{R} \mid \mathbf{x}\Box\mathbf{z}^{\tau} \in \mathcal{E}\big\}. \tag{2}$$

Axiom 1 implies $P_{\mathcal{E}}(\mathbf{1}) = 0$, and $\lim_{\mathbf{x}\to 0} P(\mathbf{x}) = +\infty$. Conversely, given a Poverty exit index, $P(\cdot)$, one defines the Poverty exit set, $\mathcal{E}_P$, as

$$\mathcal{E}_P := \big\{\mathbf{x} \in \mathbb{R}^{KN}_{++} \mid P(\mathbf{x}) \leq 0\big\}. \tag{3}$$

We now state several properties for P. As we shall see, they can be deduced from Axioms 1-4 on $\mathcal{E}$ via (2). Conversely, Axioms 1-4 can be deduced from the following properties of P, via (3).

Axiom HI (Homothetic invariance) $\forall \mathbf{x} \in \mathbb{R}^{KN}_{++}, \alpha \in \mathbb{R},\ P(\mathbf{x}\Box\mathbf{z}^{\alpha}) = P(\mathbf{x}) - \alpha$.

An easy consequence of Axiom HI is that $P(\mathbf{x}\Box\mathbf{z}^{P(\mathbf{x})}) = 0$ —which is consistent with the interpretation of $P(\cdot)$ given by (2). It follows that

$$\forall \mathbf{x},\ \mathbf{x}\Box\mathbf{z}^{P(\mathbf{x})} \in \mathcal{E}. \tag{4}$$

Axiom S (Sub-multiplicativity) $\forall \mathbf{x}, \mathbf{y},\ P(\mathbf{x}\Box\mathbf{y}) \leq P(\mathbf{x}) + P(\mathbf{y})$.

Axiom S can be interpreted as saying: averaging (in the multiplicative sense) two populations does not magnify the extent of poverty (i.e., the share of the poor within the global population), nor its intensity (i.e., the individual deprivation suffered from each individual) above the sum of the indices of the subpopulations. Since x and y are vectors of the same dimension, they correspond to populations of the same size. Hence, this axiom is a weak version of the subgroup additive decomposability applied to populations of equal size.

Axiom PH (Positive Homogeneity) : $\forall \mathbf{x} \in \mathbb{R}^{KN}_{++}, \forall \lambda \geq 0,\ P(\mathbf{x}^{\lambda}) \leq \lambda P(\mathbf{x})$.

Axiom S implies that $P(\mathbf{x}^n) \leq nP(\mathbf{x})$ for every x and every integer n. Axiom PH extends this property to any nonnegative number.

Axiom M (Monotonicity): $P(x) \leq P(y)\ \forall y \leq x$.

The aim of the last axiom is to avoid trivial indices that would be constant.

By analogy with Artzner et al. (1998), a Poverty index that satisfies Axioms HI, S, PH, M and NT is said *coherent*.

Quite similarly to the anonymity axiom for $\mathcal{E}$, the next one is not needed for the characterization of coherent poverty measures, but will prove useful.

Axiom A (Anonymity): Let $x = (x_1, .., x_n) \in \mathbb{R}^{KN}_{++}$ and $\sigma(x) = (x_{\sigma 1}, ..., x_{\sigma n}) \in \mathbb{R}^{KN}_{++}$ the vector obtained after having permuted its individual through the permutation $\sigma \in \mathcal{S}_n$. Then, $P(x) = P(\sigma(x))$.

Proposition 3.1 *(i) If a Poverty index, $P(\cdot)$, is coherent, then its Poverty exit set, $\mathcal{E}_P$, defined by 3, verifies Axioms 1-4 and is closed. Moreover, $P(\cdot) = P_{\mathcal{E}_P}(\cdot)$.*

(ii) Conversely, if a set $\mathcal{F}$ satisfies Axioms 1-4, then $P_\mathcal{F}$ is coherent, and $\mathcal{E}_{P_\mathcal{F}} = \overline{\mathcal{F}}$.[8]

(iii) $\mathcal{E}$ verifies the Anonymity axiom if, and only if, P does.

Proof.

(i) 1) $P_\mathcal{E}(\mathbf{1}) = 0$ and Monotonicity imply that $\mathcal{E}$ verifies Axiom 1.

2) If $x << \mathbf{1}$, Monotonicity implies $P_\mathcal{E}(\mathbf{x}) \geq 0$. However, we can find $\alpha > 0$ such that $x\Box\mathbf{z}^\alpha << 0$, so that $P_\mathcal{E}(x\Box\mathbf{z}^\alpha) \geq 0$. HI then implies that $\alpha \leq 0$. Contradiction. Thus, $\mathcal{E}_P$ verifies Axiom 2.

3) Axioms S and PH imply that $\mathcal{E}_P$ is multiplicatively convex.

4) If $x \in \mathcal{E}_P$, one has: $P(x^\lambda) \leq \lambda P(x) \leq 0$ for all $\lambda \geq 1$. Consequently, $\mathcal{E}_P$ is a multiplicative cone.

5) Axioms PH and S imply that the function $x \ni \mathbb{R}^{KN}_{++} \mapsto P(\exp(x))$ is convex, hence continuous. Consequently, $x \mapsto P(x)$ itself must be continuous, so that $\mathcal{E}_P$ is closed.

(ii) 0) Axioms 2 and 3 ensure that $P_\mathcal{F}$ is well-defined.

1) $\inf\{\tau \in \mathbb{R} \mid x\Box\mathbf{z}^{t+\alpha} \in \mathcal{E}\} = \inf\{\tau \in \mathbb{R} \mid x\Box\mathbf{z}^{t} \in \mathcal{E}\} - \alpha$, which proves HI.

2) Suppose that $x\Box\mathbf{z}^\lambda$ and $y\Box\mathbf{z}^\beta$ both belong to $\mathcal{E}$. Axiom 3 implies that $\big(x\Box\mathbf{z}^\lambda\big)^{\frac{1}{\alpha}}$ and $\big(y\Box\mathbf{z}^\beta\big)^{\frac{1}{1-\alpha}}$ also belong to $\mathcal{E}$ for every $\alpha \in [0,1)$. Axiom 2 then implies that $\big(x\Box y\big)\Box\mathbf{z}^{\alpha+\beta} = \big(x\Box\mathbf{z}^\lambda\big)\Box\big(y\Box\mathbf{z}^\beta\big) \in \mathcal{E}$. This proves the multiplicative convexity.

3) Suppose $x \leq y$ and $x\Box\mathbf{z}^\lambda \in \mathcal{E}$. Then, $y\Box\mathbf{z}^\lambda \geq x\Box\mathbf{z}^\lambda$, so that, by Axiom 1, $y\Box\mathbf{z}^\lambda \in \mathcal{E}$. The monotonicity of P follows.

4) If $m \geq P_\mathcal{E}(x)$, then, $x\Box z^m \in \mathcal{E}$, hence, $\forall\lambda > 0, \quad x^\lambda\Box z^{\lambda m} = \big(x\Box z^m\big)^\lambda \in \mathcal{E}$. Therefore, $P_\mathcal{E}(x^\lambda) \leq \lambda m$.

5) $\forall x \in \mathcal{F}, \quad P(x) \leq 0$. Thus, $\mathcal{F} \subset \mathcal{E}_{P_\mathcal{F}}$.

$\square$

4 Properties of coherent multidimensional poverty measures

4.1 A representation theorem and ordinality

We now provide a full characterization of the whole family of coherent Poverty exit indices. For this purpose, let us define a weighted geometric average. Given any vector in the unit simplex, $\pi \in \Delta^{KN}_+ := \{p \in \mathbb{R}^{KN}_+ \mid \sum_{k,j} p_{k,j} = 1\}$, the π-geometric average, $G^\pi(\cdot)$, is defined by:

$$G^\pi(x) := \prod_{k,h} x_{kh}^{\pi_{k,h}}.$$

[8] $\overline{\mathcal{F}}$ is the topological closure of $\mathcal{F}$.

Proposition 4.1 *The index P is coherent if, and only if, there exists a family, $\mathcal{P} \subset \Delta_+^{KN}$, of weight vectors, such that*

$$P(\mathbf{x}) = -\inf\Big\{\frac{\ln\big(G^\pi(\mathbf{x})\big)}{\ln\big(G^\pi(\mathbf{z})\big)} \mid \pi \in \mathcal{P}\Big\}.$$

Proof.

The "if" part is immediate. The "only if" part can be deduced from Proposition 2.1 in Huber (1981), and can be stated as a consequence of the bipolar theorem in linear duality theory. Consider the set

$$C := \{x \in \mathbb{R}^{KN} \mid x_{hk} = \ln(\mathbf{y}_{hk}) \text{ for some } \mathbf{y} \in \mathcal{E}\}.$$

It follows from Axiom 3 and 4 together withe the closedness of $\mathcal{E}$ that C is a convex and closed cone in $\mathbb{R}^{KN}$. Thus, its polar cone

$$C^\circ := \{\alpha \in \mathbb{R}_+^{KN} \mid \sum_{hk} \alpha_{hk} x_{hk} \geq 0 \; \forall x \in C\}$$

is also a convex and closed cone in $\mathbb{R}_+^{KN}$. The bipolar theorem implies that

$$C = \{x \in \mathbb{R}^{KN} \mid \sum_{hk} \alpha_{hk} x_{hk} \geq 0 \; \forall \pi \in \mathcal{P}\},$$

where $\mathcal{P} := \Delta_+^{KN} \cap C^\circ$. We deduce from (4) that $\ln \mathbf{x} + P(\mathbf{x}) \ln \mathbf{z} \in C$, for every $\mathbf{x} \in \mathbb{R}_{++}^{KN}$. Thus, $\forall \pi \in \mathcal{P}$, $\sum_{h,k} \pi_{hk}(\ln \mathbf{x}_{hk} + P(\mathbf{x}) \ln \mathbf{z}_{hk}) \geq 0$. Therefore,

$$P(\mathbf{x}) \sum_{h,k} \pi_{h,k} \ln \mathbf{z}_{hk} \geq -\sum_{h,k} \alpha_{hk} \ln \mathbf{x}_{hk} \quad \forall \pi \in \mathcal{P}.$$

Hence,

$$P(\mathbf{x}) \geq \sup_\pi -\frac{\ln\Big(\prod_{hk} \mathbf{x}_{hk}^{\pi_{hk}}\Big)}{\ln\Big(\prod_{hk} \mathbf{z}_{hk}^{\pi_{hk}}\Big)} = -\inf\Big\{\frac{\ln\big(G^\pi(\mathbf{x})\big)}{\ln\big(G^\pi(\mathbf{z})\big)} \mid \pi \in \mathcal{P}\Big\}.$$

Conversely, we deduce from Axiom 2 that $\ln \mathbf{x} + P(\mathbf{x}) \ln \mathbf{z} + \ln \varepsilon \notin C$ for every $\mathbf{x} \in \mathbb{R}_{++}^{KN}$ and every $0 \ll \varepsilon \ll \mathbf{1}$. Therefore, $\forall \pi \in \mathcal{P}$, $\sum_{h,k} \pi_{hk}(\ln \mathbf{x}_{hk} + P(\mathbf{x}) \ln \mathbf{z}_{hk} + \ln \varepsilon) < 0$. It follows that

$$P(\mathbf{x} \Box \varepsilon) < -\inf\Big\{\frac{\ln\big(G^\pi(\mathbf{x})\big)}{\ln\big(G^\pi(\mathbf{z})\big)} \mid \pi \in \mathcal{P}\Big\}.$$

The equality follows by continuity of $P(\cdot)$.

$\square$

Examples

a) The "utilitarian case" corresponds to $\mathcal{P} = \{((1/KN, ..., 1/KN)\}$.

b) The "Rawlsian" case corresponds to $\mathcal{P} = \Delta_+^{KN}$.

The next figure provides an illustration of the typical geometry of $\mathcal{E}$.

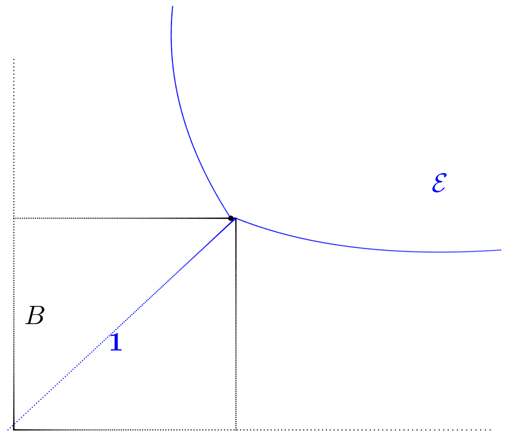

Fig 1. A piecewise smooth poverty exit set

Observe that, in general, the frontier of the set $\mathcal{E}$ need not be smooth, as there is typically a kink at $\mathbf{1}$. The "utilitarian" case corresponds to the situation where the two branches of hyperbola coincide with the hypercurve : $\prod_{h,k} \mathbf{x}_{h,k} = \text{cst}$. It is the unique case where the boundary of $\mathcal{E}$ is a smooth submanifold. The larger the set $\mathcal{P}$, the smaller the subset $\mathcal{E}$. Finally, the Rawlsian case corresponds with the situation where $\mathcal{E}$ coincides with the affine nonnegative orthant

$$\mathcal{E} = \mathbf{1} + \mathbb{R}_+^{KN}.$$

Notice that weights in $\mathcal{P}$ can differ across both individuals and dimensions. When P (or, equivalently, $\mathcal{E}$) verifies Anonymity, the set of weights, $\mathcal{P}$, reduces to weights over dimensions. The weighted geometric average now becomes:

$$G^{\tilde{\hat{\pi}}}(x) := \prod_{k,h} x_{kh}^{\frac{\hat{\pi}_k}{N}} \quad \forall \hat{\pi} \in \hat{\mathcal{P}} \subset \Delta_+^K.$$

Corollary 4.1 *The index P is coherent and anonymous if, and only if, there exists a family, $\hat{\mathcal{P}} \subset \Delta_+^K$, of weights over dimensions such that*

$$P(\mathbf{x}) = -\inf\Big\{ \frac{\ln\big(G^{\pi}(\mathbf{x})\big)}{\ln\big(G^{\pi}(\mathbf{z})\big)} \mid \pi \in \hat{\mathcal{P}} \Big\}.$$

Thanks to Theorem 4.1, whether it is anonymous or not, a coherent index P can also easily be shown to be *ordinal* in the following sense.

Ordinality. A measure, Q, is said to be ordinal if the following holds. Given some diagonal matrix $\Lambda = (\lambda_{jj})_{j=1,\dots,KN}$ with positive entries ($\lambda_{jj} > 0$), given also a social status matrix $y \in \mathcal{M}_{N\times K}(\mathbb{R}_{++})$, and a cut-off vector, $\mathbf{z} \in \mathbb{R}_{++}^{KN}$, one has:

$$Q(y; \mathbf{z}) = Q(y\Lambda; \mathbf{z}\Lambda).$$

An example will easily illustrate how this abstract property solves most of the problems related to ordinal data. Consider the question: "Which kind of toilet facility does your household have ?", together with three possible answers:

a. "Open defacation field".

b. "Shared flush".

c. "Private flush".

Of course, the metric between each one of these answers does not have any sensible meaning. To circumvent this issue, it suffices to capture this question through two dimensions, each of them accepting two answers, $\{a, b\}$ and $\{a, c\}$, each captured by two variables $\{0, \alpha\} \subset \mathbb{R}$ and $\{0, \beta\} \subset \mathbb{R}$ respectively. Ordinality then ensures that the choice of α and β does not matter.

Going back to coherent poverty measures, it is straightforward that, for any $x \in \mathbb{R}^{KN}_{++}$ and any Λ as above, $x\Lambda/\mathbf{z}\Lambda = x/\mathbf{z}$. Thus, as we only deal with normalized achievements, any Multidimensional Poverty Index is ordinal.

4.2 Who is poor ?

In this subsection, we confine ourselves to the subfamily of anonymous coherent Poverty indices. Consequently, P is associated with a set, $\mathcal{P} \subset \Delta^k_+$, of K-dimensional vector of weights, $\pi = (\pi_k)_k$ –one for each dimension–, belonging to the unit simplex.

We now provide an answer to the question: "who is poor" ? Regarding this issue, two kinds of approach have been explored in the literature.[9] The "union" approach regards a person who is deprived in one dimension as being poor at the multidimensional level. This is usually acknowledged to be overly inclusive and lead to exaggerate estimates of poverty. By contrast, the "intersection" approach requires a person to be deprived in all dimensions before getting considered as poor. This is often considered too constricting, and may lead to untenably low estimates of poverty. We now show that the natural definition of a poor person that follows from our approach leads to an endogenous determination that is always strictly less inclusive than the "union" approach and weakly more inclusive than the "intersection" approach. Therefore it lies somewhere between these two extremes, and in fact, it turns out that only the Rawlsian case coincides with the "intersection" viewpoint.

Two examples will help identify how the determination of poor persons occurs in our setting. Consider the case where $N = 1$, i.e., the population consists of a single person. Then, clearly, this single person, i, will be poor whenever the population is so, i.e., when $P(\mathbf{x}_i) < 0$. Next, suppose that the population is made of n identical people. Again, each person will be poor if the population is so, i.e., if, and only if, $\prod_k x_i^{\pi_k} < \prod_k \mathbf{z}_i^{\pi_k}$ for every $\pi \in \mathcal{P}$.[10]

It is this latter condition that we adopt as a definition. Indeed, a simple continuity argument explains why now other choice can be made: Take $0 \ll \varepsilon \ll \mathbf{1}$; one has $x\Box\varepsilon$ poor and $G^\pi(\mathbf{x}) < 1$ for any π. However, $\lim_{\varepsilon \to \mathbf{1}} G^\pi(\mathbf{x}) = 1^-$. Thus, no population such that $G^\pi(\mathbf{x}) < 1$ can be considered as non-poor.

Definition 4.1 Given a coherent Poverty index, P, associated with a set $\mathcal{P} \subset \Delta^{KN}_+$ of weights, a person, i, is poor whenever

$$\prod_k \mathbf{x}_{i,k}^{\pi_k} < \prod_k \mathbf{z}_k^{\pi_k} \; \forall \pi \in \mathcal{P}$$

or, equivalently, when

[9] See, e.g., Alkire and Foster (2008) and Villar (2010).

[10] Notice that, here, x is *not* normalized.

$$\sup_{\pi \in \mathcal{P}} G^{\pi}(\mathbf{x}_i) < 1.$$

In the "utilitarian" case (where $\mathcal{P}$ reduces to the uniform singleton), this definition coincides with the one introduced by Villar (2010).

As an illustration, let us consider a society with two dimensions. The poor are all the individual strictly below the two branches of hyperbola:

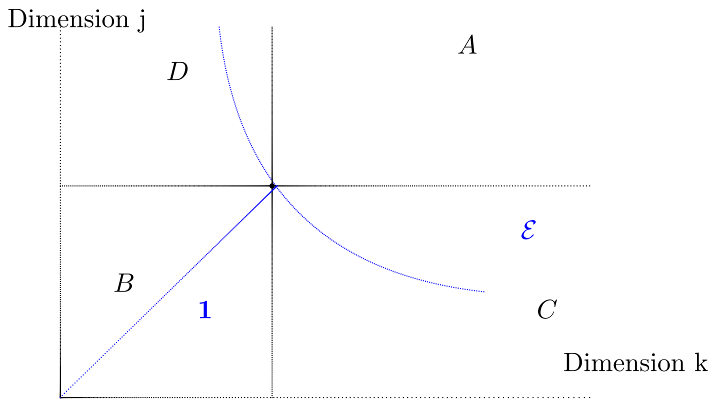

Fig 2. Who is poor ?

The set $\mathcal{E}$ is always larger than the one defined by the intersection approach, and is always contained in the one provided by the union approach. The Rawlsian case, here, coincides with the intersection approach.

4.3 Other properties

Here are the properties verified by coherent Poverty indices. When they are evident, proofs are left to the reader.

1. Multiplicative decomposability : Suppose that x_1 (resp. x_2) is a population of size n_1 (resp. n_2). Let us denote by $\langle x_1, x_2 \rangle$ the population of size $n = n_1 + n_2$, obtained by merging the first two. One has:

$$G^{\pi}(\langle x_1, x_2 \rangle) = \left[G^{\pi}(x_1)\right]^{\frac{n_1}{n}} \left[G^{\pi}(x_2)\right]^{\frac{n_2}{n}} \ \forall \pi \in \mathcal{P},$$

so that

$$P\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = P(\mathbf{x}_1^{\frac{n_1}{n}} \Box \mathbf{x}_2^{\frac{n_2}{n}}).$$

2. The next property is a special case of multiplicative decomposability:

Replication invariance : For any population, x,

$$P\langle \mathbf{x}, \mathbf{x} \rangle = P(\mathbf{x}).$$

3. Path independence: One can aggregate individual unidimensional values first across dimensions and then across agents, or viceversa, obtaining the same result.

4. The next property asks that a reduction of size $\delta > 0$ in the deprivation with respect to dimension k of a poor person i who is worse off in this dimension than another poor person, j, more than compensates an increase of the same size, δ, in the deprivation of j, provided their relative positions remain unaltered. Formally, if $x_{jk} - x_{ik} \geq 2\delta$, $y_{ik} = x_{ik} + \delta$ and $y_{jk} = x_{jk} - \delta$, while $y_{h\ell} = x_{h\ell}$ $\forall (h, \ell) \notin \{(i, k), (j, k)\}$.

Transfer principle. $P(\mathbf{y}) \leq P(\mathbf{x})$.
Indeed,

$$x_{ik} x_{jk} < (x_{ik} + \delta)(x_{jk} - \delta) = y_{ik} y_{jk}.$$

It follows that $\left[x_{ik} x_{jk}\right]^{\alpha_k} \leq \left[y_{ik} y_{jk}\right]^{\alpha_k}$, for every $\alpha_k \geq 0$. The conclusion follows.
Also observe that the geometric mean is a distribution sensitive measure that penalizes the dispersion of the individual values, relative to the arithmetic mean. In particular, for two distributions with identical mean values it assigns higher value of the intensity of the poverty to that in which the distribution of the yij values is more disperse.

5. The reduction in the deprivation of dimension k required to compensate an increase in the deprivation of dimension ℓ is smaller the smaller the initial level of achievement in ℓ. This feature simply follows from the decreasing marginal rate of substitution of the individual poverty index across achievement dimensions). Obviously, this property cannot be satisfied by any (weighted) arithmetic measure.

6. The Poverty focus requirement says that only changes within the population, $N(y; \mathbf{z})$, of poor affect P. This property is not fulfilled, in general, by coherent indices as these capture some kind of substitutability among poor and non-poor. However, as long as the cut-off, $\mathbf{z}$, is exogenous, one easy way to recover Poverty focus consists in censoring achievements as follows, before normalizing them:[11]

$$\tilde{x}_{ik} := \begin{cases} x_{ik} & \text{if } x_{ik} < \mathbf{z}_{ik} \\ \mathbf{z}_{ik} & \text{if } x_{ik} \geq \mathbf{z}_{ik}. \end{cases}$$

7. The same censoring provides us with the Deprivation focus property, namely: only changes in dimensions where poor people are deprived affect P.

8. Following Kolm (1977) and Alkire and Foster (2008), we can check how much a coherent poverty index, P, is sensitive to inequality in the distribution of achievements and deprivations. There are several ways to do this. One way consists in considering mean-preserving spreads, i.e., transformations of a given population that increase the spreads of the achievements with respect to their arithmetic mean without affecting the mean itself. (Such transformations are the reversal of the change considered above for the Transfer principle.) An inequality-sensitive Index should be decreasing with respect to such transformations. Formally, an increase of size $\delta > 0$ in the deprivation with respect to dimension k of i should not compensate an decrease of the same size, δ, in the deprivation of j. Formally, if $x_{jk} - x_{ik} \geq 0$, $y_{ik} = x_{ik} - \delta$ and $y_{jk} = x_{jk} + \delta$, while $y_{h\ell} = x_{h\ell}$ $\forall (h, \ell) \notin \{(i, k), (j, k)\}$, then, by the same argument as for the Transfer principle, we get:

[11]This is standard practice, see Alkire and Foster (2008).

Mean-preserving spread sensitivity $P(\mathbf{x}) < P(\mathbf{y})$.

9. Following Atkinson and Bourguignon (1982) and Alkire and Foster (2008), we say that x is obtained from y by a *simple rearrangement among the poor* if the achievements of two poor persons, i and j, have been reallocated so that, for each dimension k:

$$(x_{ik}, x_{jk}) = (y_{jk}, y_{ik}) \quad \text{ or } (x_{ik}, x_{jk}) = (y_{ik}, y_{jk}),$$

while the achievements of anyone else remain untouched. If, in addition, y_i and y_j are comparable but x_i and x_j are not, then x is said to be obtained from y by an *association decreasing rearrangement among the poor*. Reducing inequality this way does trivially decrease any coherent multidimensional Poverty Index:

$$P(\mathbf{y}) \leq P(\mathbf{x}).$$

This property is called Weak arrangement.

10. Another way to test the sensitivity towards inequality of an Index consists in averaging the achievement vectors, y_i and y_j of two poor persons, i and j in such a way that i now exhibits $x_i := (1 - \lambda)y_i + \lambda y_j$ (with $\lambda \in (0, 1)$) and $x_j := \lambda y_i + (1 - \lambda)y_j$. The new population (x_i, x_j) is viewed as being unambiguously less unequal than the original one, (y_i, y_j), which should result in a lower or equal value of the multidimensional poverty index. Here, we translate linear convex combinations in geometric combinations, so as to arrive at the following definition. We say that $x \in \mathcal{M}_{n\times k}(\mathbb{R}_{++})$ is obtained from $y \in \mathcal{M}_{n\times k}(\mathbb{R}_{++})$ by a *geometric averaging of achievements among the poor* if, for every poor i, there exist weights $(\alpha_j)_{j\in N(y;\mathbf{z})} \in \Delta_+^{n(y;\mathbf{z})}$ such that

$$x_i = \prod_{j\in N(y;\mathbf{z})} y_j^{\alpha_j},$$

while non poor persons are not affected (i.e., $x_i = y_i$ for $i \notin N(y; \mathbf{z})$).

Multiplicative weak transfer. If x is obtained from y by a geometric averaging of achievements among the poor, then one should have $P(\mathbf{x}) \leq P(\mathbf{y})$.

However, this property is not satisfied by a coherent measure, in general. Consider, for example, a population , (a, b), consisting in 2 persons and a single dimension (with $a < b < 1$). The population $(a^{1/3}b^{2/3}, b)$ is obtained from (a, b) by a geometric averaging of achievements among the poor, but:

$$G(a^{1/3}b^{2/3}, b) > G(ab).$$

References

Alkire, S. and Foster, J. (2008). Counting and multidimensional poverty measurement. *OPHI Working Paper*. 2, 3, 12, 14, 15

Artzner, P., Delbaen, F., Ebera, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228. 2

Atkinson, A. and Bourguignon, F. (1982). The comparison of multidimensional distribution of economic status. *Review of Economic Studies*, 49(2):183–201. 15

Huber, P. (1981). *Robust Statistics*. Wiley, New York. 10

Kolm, S. (1977). Multidimensional egalitarianism. *Quarterly Journal of Economics*, 91:1–13. 14

Seth, S. (2009). A class of association sensitive multidimensional welfare indices. *Journal of Economic Inequality (forthcoming)*. 4

Villar, A. (2010). A new approach to multidimensional poverty measurement. *Pablo Olavide university Working Paper*. 1, 2, 3, 12, 13